1974

# The generation of Poisson random deviates

Pak, Chae Ha

Monterey, California : Naval Postgraduate School

http://hdl.handle.net/10945/17038

# THE GENERATION OF POISSON RANDOM DEVIATES

Chae Ha Pak

# NAVAL POSTGRADUATE SCHOOL
## Monterey, California

# THESIS

THE GENERATION OF POISSON
RANDOM DEVIATES

by

Chae Ha Pak

September 1974

Thesis Advisor: Peter A. W. Lewis

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) The Generation of Poisson Random Deviates | | 5. TYPE OF REPORT & PERIOD COVERED Master's Thesis September 1974 |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Chae Ha Pak | | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940 | | 12. REPORT DATE September 1974 |
| | | 13. NUMBER OF PAGES 34 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940 | | 15. SECURITY CLASS. (of this report) Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Poisson Random Distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Three approximation methods for generating outcomes on Poisson random variables are discussed. A comparison is made to determine which method requires the least computer execution time and to determine which is the most robust approximation. Results of the comparison study suggest the



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block #20 continued

method to choose for the generating procedure depends on the mean value of Poisson random variable which is being generated.

# The Generation of Poisson Random Deviates

by

Chae Ha Pak
Lieutenant, Korean Navy
B.S., Korean Naval Academy, 1969

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1974

ABSTRACT

Three approximation methods for generating outcomes on Poisson random variables are discussed. A comparison is made to determine which method requires the least computer execution time and to determine which is the most robust approximation. Results of the comparison study suggest the method to choose for the generating procedure depends on the mean value of Poisson random variable which is being generated.

TABLE OF CONTENTS

LIST OF TABLES

## LIST OF FIGURES

# I. INTRODUCTION

It is frequently desired to generate Poisson random variates in simulations. There are standard exact methods for doing this; the problem arises when a computer is used to generate the Poisson random number which has a large mean. For example, generating one random number such as 105 from a Poisson distribution with mean 100 needs at least 105 calls to a pseudo-random number (uniform (0,1)) generator. Computer time requirements become important cost factors when considering various methods for generating random numbers.

The objective of this paper is to examine several approximated ways of generating Poisson random variables and to determine the method which gives minimum execution time and small mean absolute deviation according to the Poisson mean value. The mean value is the only parameter in the distribution. Comparison statistics to determine the best approximation to the Poisson distribution are the cumulative probability, mean absolute deviation and Kolmogorov-Smirnov test. Finally a composite generating procedure according to the mean value is suggested.

The following notation is used in this study.

$U_i$ denotes a uniform (0,1) random variable;

N denotes a Poisson random variable where mean is m;

Z denotes a random variable from a unit normal distribution.

## II. GENERATION OF POISSON DISTRIBUTED VARIATES

### A. THE POISSON DISTRIBUTION

A random variable N with integer values has a Poisson distribution if

$$\text{prob}\ \{N = m\} = \frac{e^{-m} m^{n}}{n!} \qquad n = 0,1,2\ldots \quad .$$

In order to generate a Poisson random number N from a Poisson distribution with mean m, the following algorithm is presented. It is the standard exact method for generating these deviates.

Let $U_i$, $i = 1,2,\ldots,$ be independent uniform (0,1) random variates. The Poisson variate, N, is computed as:

$$N = \begin{cases} 0 & \text{if } U_i \leq e^{-m} \\ k & \text{if } \prod_{i=1}^{k} U_i > e^{-m} \geq \prod_{i=1}^{k+1} U_i \end{cases}$$

where N is non distributed as a Poisson with mean m, i.e., Prob $(N=n) = e^{-m}m^{n}/n!$ .

Equivalently, since the logarithm is a monotone transformation, we have

$$N = \begin{cases} 0 & \text{if } \ln(U_i) \leq \ln(e^{-m}) = -m \\ k & \text{if } \ln(\prod_{i=1}^{k} U_i) = \sum_{i=1}^{k} \ln(U_i) > \ln(e^{-m}) \end{cases}$$

$$= -m \geq \sum_{i=1}^{k+1} \ln(U_i) = \ln\ (\prod_{i=1}^{k+1} U_i).$$

Changing the signs and direction of the inequalities we have:

$$N = \begin{cases} 0 & \text{if } -\ln(U_i) \geq m \\ k & \text{if } \sum_{i=1}^{k} - \ln(U_i) < m \leq \sum_{i-1}^{k+1} - \ln(U_i) \end{cases}$$

or letting $E_i = - \ln(U_i)$

$$N = \begin{cases} 0 & \text{if } E_i \geq m \\ k & \text{if } \sum_{i=1}^{k} E_i < m \leq \sum_{i=1}^{k+1} E_i \end{cases}$$

where the $E_i$ are exponentially distributed variates with mean 1.

If n multiplications of uniform (0,1) random numbers is strictly greater than $e^{-m}$ and if n+1 multiplication of uniform (0,1) random number is equal and less than $e^{-m}$ then n is the Poisson random number. Generally generating one random number from Poisson process with parameter m requires on the average m+1 uniform (0,1) random numbers. This is because the number generated is n+1. When m is large it is clear that generating Poisson random numbers with the above method, although it is exact takes a lot of computer time and this method may be uneconomical. In addition, the large number of multiplications can produce serious precision problems on a digital computer.

## B. APPROXIMATIONS FOR POISSON VARIATES

### 1. Normal Approximation

In a Poisson process with parameter $\lambda$ it is necessary to generate random variables from the Poisson distribution with parameter (mean) m. Now look at counts in $(0,x)$, where x satisfies $\lambda x = m$. The central limit theorem says that as m goes to positive infinity, (or when x gives the infinity in Poisson process with fixed $\lambda$), then N, which has mean m and variance m, is such that $N+\frac{1}{2}-m/m^{\frac{1}{2}}$ is approximately distributed as a unit normal random variable. Denote a random variable from a unit normal distribution by Z. So N is distributed approximately as $m^{\frac{1}{2}}Z + m$. In order to generate Poisson random numbers from the normal distribution, first generate Z; then let

$$\tilde{N} = \begin{cases} 0 & \text{if } m^{\frac{1}{2}}Z + m - 0.5 < 1 \\ \lfloor m^{\frac{1}{2}}Z + m - 0.5 \rfloor & \text{otherwise} \end{cases}$$

where $\lfloor a \rfloor$ denotes the greatest integer less than or equal to a; also known as the "floor" of a. $\tilde{N}$ is then the approximated Poisson random variate.

### 2. Square Root Transformation of Poisson Distribution

If N is a Poisson random variable with mean m then $Y = \sqrt{N + 3/8}$ is approximately distributed as a normal distribution with mean $m^{\frac{1}{2}}$ and variance $\frac{1}{4}$. This result is due to Bartlett [Ref. 9].

This method is derived as follows: let $Y = \sqrt{N + C} \sim N(\mu,\sigma^2)$ where C is a non-negative constant. Let $t = N - m$ and $m' = m + C$. Define coefficients for $s = 1,2,3,\ldots$ by

$$a_s = (-1)^{s+1} \frac{1\cdot(-1)\cdot(-3)\cdots(-2s+3)}{2^s\cdot s!}$$

Then for any $t \geq - m'$ we have a Taylor series expansion.

$$Y = \sqrt{m'}\ \{1+A_1 \frac{t}{m'} - A_2(\frac{t}{m'})^2 +\cdots+(-1)^s A_{s-1}(\frac{t}{m'})^{s+1}\}$$

$$+ R_s.$$

If $t > 0$, we see at once $|R_s| < A_s t^s/(m')^{s-\frac{1}{2}}$ converges and is bounded (F. J. Anscombe (4)). We note now that the moments of t are $\mu_1 = 0$, $\mu_2 = m$, $\mu_3 = m$, $\mu_4 = 3m^2 + m,\ldots$, which give

$$\text{Var}\ (Y) \sim \tfrac{1}{4}\ (1 + \frac{3-8C}{8m} + \frac{32C^2 - 52C + 17}{32m^2}),$$

so that when $C = 3/8$ Var $(Y) \sim \frac{1}{4}\ (1 + 1/16m^2)$. Also

$$E(Y) \sim \sqrt{m+C} - \frac{1}{8m^{\frac{1}{2}}} + \frac{24C - 7}{128m^{3/2}}.$$

Let $XNR = \sqrt{N + 3/8}$. Then XNR is approximately normally distributed with mean $\sqrt{m}$ and variance $\frac{1}{4}$.

$$Z = \frac{XNR - \sqrt{m + 3/8}}{\frac{1}{2}},$$

$$XNR = \frac{Z}{2} + \sqrt{m + 3/8},$$

$$\sqrt{N + 3/8} = \frac{Z}{2} + \sqrt{m + 3/8},$$

thus set

$$\tilde{N} = \begin{cases} 0 & \text{if } \left(\frac{Z}{2} + \sqrt{m+3/8}\right)^2 - 3/8 < 1 \\ \left\lfloor \left(\frac{Z}{2} + \sqrt{m+3/8}\right)^2 - 3/8 \right\rfloor & \text{otherwise} \end{cases}$$

$\tilde{N}$ is then the approximated Poisson variate.

We now need to calculate the probability distribution of $\tilde{N}$ obtained in this way from the square root transformation. We want the probability that

$$n - 1 + 3/8 < XNR^2 < n + 3/8$$

if we divide by the variance $\frac{1}{4}$ we get

$$4(n - 1 + 3/8) < XNR^2\ 4 < 4(n + 3/8).$$

Note that $4 \cdot XNR^2$ is distributed with a non-central $\chi_1^2$ distribution. The non-central $\chi_1^2$ density is

$$f_x(x) = \frac{e^{-\frac{1}{2}(x+\mu^2/\sigma^2)}}{2\sqrt{x}\ \ \sqrt{2\pi}} \left[e^{-\frac{\mu}{\sigma}\sqrt{x}} + e^{\frac{\mu}{\sigma}\sqrt{x}}\right].$$

Thus if $\mu=m^{\frac{1}{2}}$, $\sigma^2=(\frac{1}{2})^2$, then

$$f_x(x) = \frac{e^{-\frac{1}{2}(x+m/\frac{1}{4})}}{2\sqrt{x}\ \ \sqrt{2\pi}} \left[e^{-(\sqrt{m}/\frac{1}{2})\sqrt{x}} + e^{(\sqrt{m}/\frac{1}{2})}\right]$$

and

$$\text{prob}(x=N) = \int_0^{\frac{N+3/8}{(\frac{1}{2})^2}} f_x(x)dx - \int_0^{\frac{N+3/8-1}{(\frac{1}{2})^2}} f_x(x)dx$$

This allows us to evaluate directly how well the distribution of $\tilde{N}$ approximates the distribution of a Poisson variate with parameter m.

Note that since in the LLRANDOM package it takes the same amount of time to generate 5 uniform random variables as it takes to generate a normal random variable, the procedures will be competitive timewise once m is much greater than 5.

3. Cube Root Transformation of Poisson Distribution

If N is a Poisson random variable with mean m then $Y = \sqrt[3]{N-1/24}$ is approximately distributed as a normal distribution with mean $m^{1/3}$ and variance $1/9\sqrt[3]{m}$ when $N \geq 1$. This comes from the following derivation which is essentially the same procedure as for the square root transformation. Suppose $Y = \sqrt[3]{N+C}$ is distributed as $N(\mu,\sigma^2)$. Let $t = N - m$ and $m' = m + C$. Define coefficients for $s = 1,2,3,\ldots$ by

$$a_s = \frac{(-1)^{s+1}\ 1\cdot(-2)(-5)(-8)\cdots(4-3s)}{3^s\ s!}.$$

For any $t \geq -m'$ we have the Taylor series expansion

$$Y = \sqrt[3]{m'}\ \{1 + a_1 \frac{t}{m'} - a_2(\frac{t}{m'})^2 + a_3(\frac{t}{m'})^3 - a_4(\frac{t}{m'})^4 \cdots$$

$$\cdots+(-1)^s\ a_{s-1}\ (\frac{t}{m'})^{s-1}\} + R_s$$

if $t > 0$, we see at once that $|R_s| \leq a_s t_s/(m')^{s-1/3}$ converges. Therefore,

$$R_s = \frac{f^{(s)}\ (1 + \theta \frac{1}{m'})}{s!} (\frac{t}{m})^s \quad 0 \leq \theta \leq 1$$

$$|t| \leq m'$$

$$R_s(m')^{-1/3} = (1 + \frac{t}{m'})^{1/3} - \{1 + a_1 \frac{t}{m'} - \cdots + (-1)^s\ a_{s-1}(\frac{t}{m})^{s-1}\}$$

$$\frac{R_s(m')^{-1/3}}{t^s} = \sum_{i=s}^{\infty} (-1)^{i+1}\, a_i\left(\frac{t}{m'}\right)^i$$ converges and is

bounded.

We note now that the moments of t are $\mu_1=0$, $\mu_2=m$, $\mu_3=m$, $\mu_4=3m\ +m,\ldots,$ giving us

$$E(Y) \cong \sqrt[3]{m+C} - \frac{1}{18}\frac{1}{m^{2/3}} + \ldots$$

and

$$Var(Y) \cong (\sqrt[3]{m'})^2 \left\{ \frac{1}{9}\cdot\frac{1}{(m')^2}\, xm - \frac{1}{81}\frac{2m^2+m}{(m')^4} \cdots\right\}$$

$$= \frac{1}{9\sqrt[3]{m}} - \frac{1}{\sqrt[3]{m^4}}\left( \frac{4}{27}C + \frac{1}{162} \right)\cdots \quad .$$

If $C = \frac{1}{-24}$ then

$$Var(Y) \cong \frac{1}{9\sqrt[3]{m}} \; .$$

Let

$$YNR = \sqrt[3]{N - 1/24} \qquad N \geq 1$$

$$Z = \frac{YNR - \sqrt[3]{m + 1/24}}{\frac{1}{3\sqrt[6]{m}}}$$

$$YNR = \sqrt[3]{m + 1/24} + (1/3\sqrt[6]{m})\ Z$$

$$\sqrt[3]{N-1/24} = \sqrt[3]{m + 1/24} + (1/3\sqrt[6]{m})\ Z.$$

Thus set

$$\tilde{N} = \begin{cases} 0 & \text{if } (\sqrt[3]{m+1/24}+\frac{1}{3}\sqrt[6]{m}\ Z)^3+\frac{1}{24} < 1 \\ \left\lfloor (\sqrt[3]{m+1/24} + \frac{1}{3}\sqrt[6]{m}\ Z)^3 + \frac{1}{24} \right\rfloor & \text{otherwise} \end{cases}$$

$\tilde{N}$ is now the approximated Poisson variate.

## III. EVALUATION OF THE METHODS

Generally, generating Poisson random number from the exact method is known to take a long execution time, since one generated random number "N" requires on the average N multiplications of uniform (0,1) random numbers. Therefore, generating Poisson variables with the exact method (which gives the best accuracy) is good for a small mean, m, while the approximation methods, which take shorter execution time but with less accuracy, should be used for large m. Here we need a trade-off between execution time and accuracy to choose the generation method according to the mean value of the Poisson distribution.

The comparison statistics show how closely the methods approximate the original Poisson distribution. In the Kolmogorov-Smirnov test, all approximations are accepted at significance level $\alpha = .05$. From the comparison of the empirical probability distributions and the mean absolute deviations from the exact distribution, the optimal generation procedure based on the mean value, m, is as follows:

| | Method | Mean (m) |
|---|---|---|
| 1. | exact Poisson distribution | if $0 \leq m \leq 20$ |
| 2. | square root transformation | if $20 < m \leq 100$ |
| 3. | normal approximation | if $m > 100$. |

The cube root transformation should not be adopted because it is far less accurate compared to the square root

transformation and the normal approximation. The following tables and figures show the comparison statistics of the three approximations versus the exact Poisson distribution.

Table I assesses the accuracy of the normal, square root, and cube root approximations to the exact Poisson distribution at selected points. Figure 1 illustrates graphs of the empirical cumulative distribution functions of the three approximations to that of an exact Poisson distribution with mean 10. Figure 2 shows graphs of the mean absolute deviations of the three approximations from the exact Poisson cumulative distribution function for various mean values. The mean absolute deviation is defined as:

$$\{ \frac{1}{k-1} \sum_{i=1}^{k} (P_i' - P_i)^2 \}$$

where

$P_i'$ is the CDF of the approximating distribution;

$P_i$ is the CDF of the exact Poisson distribution; and

k is the sample size.

Table II is a summary of a comparison of the empirical distributions produced by the three approximation methods to the exact Poisson distribution by means of the one sample Kolmogorov-Smirnov test. The null hypothesis is that the approximated distributions are Poisson against the alternative that they are not Poisson.

Lastly, Table III and Figure 3 analyze the sensitivity of the square root transformation to the constant C. In the

derivation of the square root transformation, C was chosen as 3/8. Different constants were used in order to find the most robust constant to use, i.e., the constant which yielded the smallest mean absolute deviation from the exact Poisson. The value $C = 13/18$ was found to be the most robust. Note that this value of C was used in the approximation when making the comparisons with the other methods.

Table I. Accuracy of the Normal, Square Root, and Cube Root Approximations to the Exact Poisson Distribution.

| Poisson mean, m | Observed Value, n | $P(Z_{(m)} \leq n)$ | | | |
|---|---|---|---|---|---|
| | | Exact | Normal Approx | Square Root Transformation C=13/18 | Cube Root Transformation |
| 5 | 2 | 0.08422 | 0.07301 | 0.07719 | 0.10491 |
| | 3 | 0.14037 | 0.11939 | 0.14051 | 0.18491 |
| | 6 | 0.14622 | 0.16036 | 0.14614 | 0.16681 |
| 15 | 10 | 0.04861 | 0.04485 | 0.04704 | 0.09045 |
| | 14 | 0.10244 | 0.09937 | 0.10288 | 0.05512 |
| | 18 | 0.07062 | 0.07622 | 0.07073 | 0.06348 |
| 20 | 14 | 0.03874 | 0.03633 | 0.03746 | 0.08819 |
| | 19 | 0.08884 | 0.08683 | 0.08891 | 0.08726 |
| | 22 | 0.07692 | 0.08050 | 0.07672 | 0.05755 |
| 40 | 36 | 0.05394 | 0.05161 | 0.05409 | 0.02552 |
| | 39 | 0.06296 | 0.06223 | 0.06307 | 0.06336 |
| | 42 | 0.05850 | 0.05995 | 0.05841 | 0.05675 |
| 61 | 55 | 0.03960 | 0.03910 | 0.03963 | 0.00662 |
| | 59 | 0.05019 | 0.04999 | 0.05020 | 0.04973 |
| 82 | 82 | 0.04407 | 0.04402 | 0.04402 | 0.03289 |
| | 87 | 0.03686 | 0.03649 | 0.03683 | 0.03801 |
| | 89 | 0.03165 | 0.03243 | 0.03149 | 0.03991 |
| 90 | 94 | 0.03775 | 0.03805 | 0.03759 | 0.0255 |
| | 97 | 0.03112 | 0.03182 | 0.03098 | 0.0435 |
| 100 | 93 | 0.03223 | 0.03241 | 0.03217 | 0.0410 |
| | 99 | 0.03997 | 0.03980 | 0.03994 | 0.03605 |
| 120 | 112 | 0.02881 | 0.02847 | 0.02866 | 0.03627 |
| | 116 | 0.03477 | 0.03438 | 0.03462 | 0.00130 |

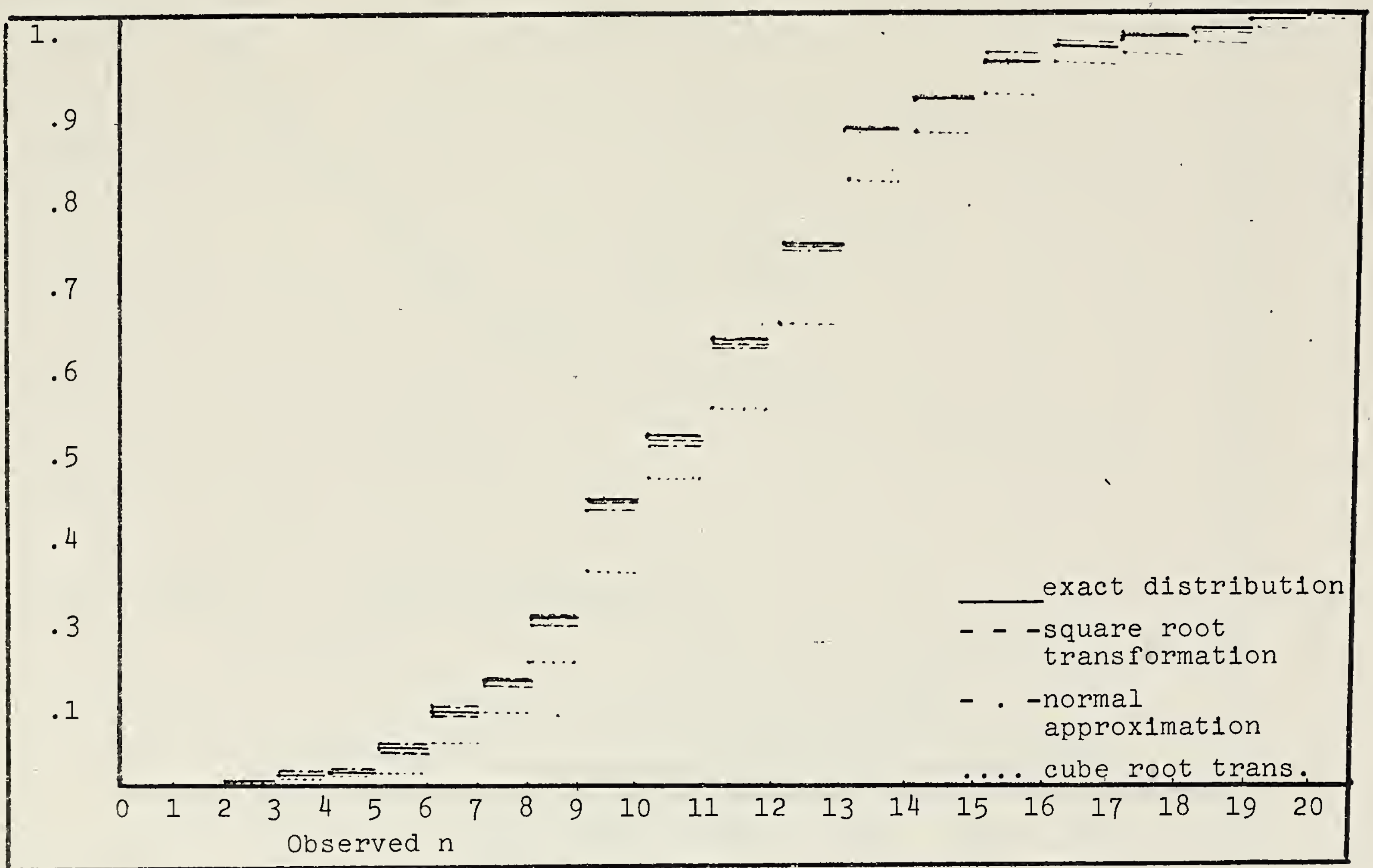


Figure 1. Comparison of Empirical Cumulative Distribution Functions for Poisson Distribution with Mean 10.

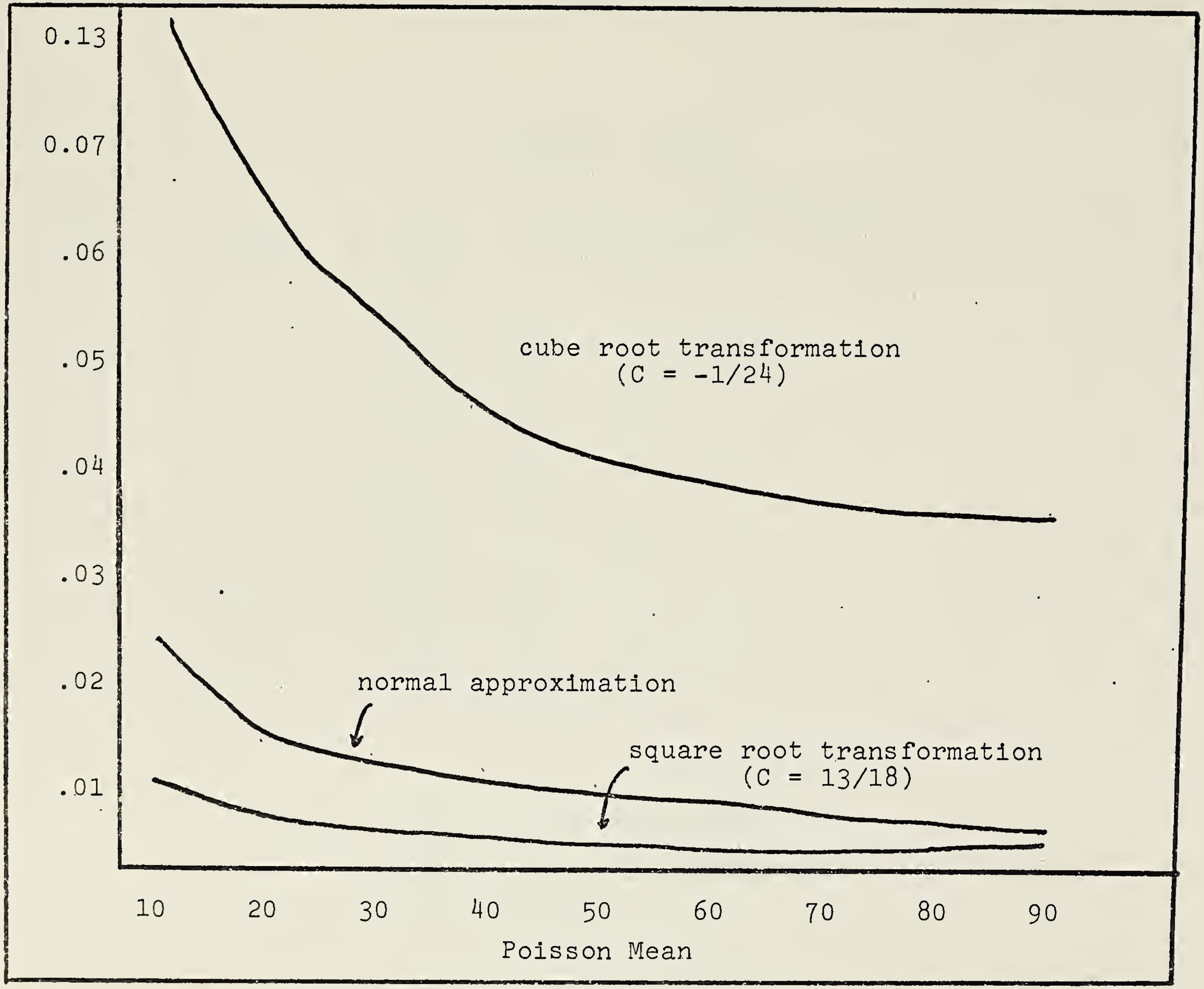


Figure 2. Mean Absolute Deviations of the Three Approximations.

| Poisson mean, m | Sample Size | $MAX\lvert\tilde{F}(x) - F(x)\rvert$ | | | Critical Value, at $\alpha = .05$ | Accept or Reject |
|---|---|---|---|---|---|---|
| | | Normal Approx | Square Root Transformation (C=13/18) | Cube Root Transformation (C=-1/24) | | |
| 10 | 19 | 0.021 | 0.010 | 0.092 | 0.312 | Accept |
| 30 | 32 | 0.012 | 0.006 | 0.052 | 0.240 | Accept |
| 50 | 40 | 0.009 | 0.005 | 0.040 | 0.215 | Accept |
| 70 | 47 | 0.008 | 0.004 | 0.034 | 0.198 | Accept |
| 90 | 52 | 0.008 | 0.005 | 0.030 | 0.189 | Accept |

Table II.

One Sample Kolmogorov-Smirnov Test. .

| C | Mean Absolute Deviation $\sqrt{\frac{1}{k-1}\sum_{i=1}^{k}(P_i' - P_i)^2}$ | |
|---|---|---|
| | Poisson Mean = 20 | Poisson Mean = 80 |
| 3/8 | 0.019 | 0.009 |
| 5/9 | 0.016 | 0.006 |
| 7/12 | 0.014 | 0.005 |
| 11/18 | 0.011 | 0.005 |
| 23/36 | 0.009 | 0.004 |
| 8/12 | 0.007 | 0.003 |
| *13/18 | 0.006 | 0.002 |
| 9/12 | 0.007 | 0.003 |
| 5/6 | 0.011 | 0.004 |

* optimal constant for C.

Table III.

Sensitivity of the Square Root Transformation to the Constant C for Poisson Distributions with Means m = 20 and m = 80.

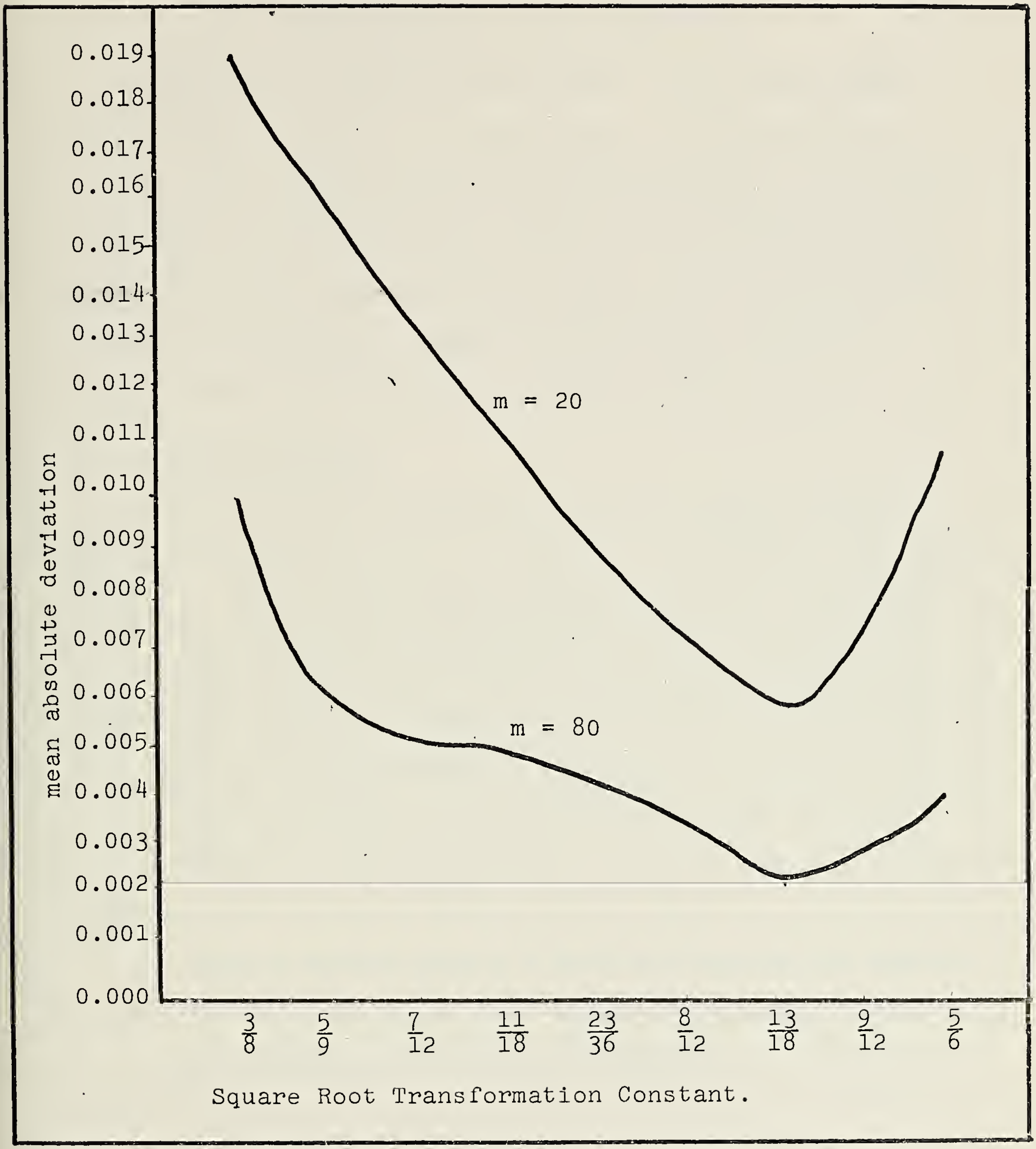


Figure 3. Plot of Sensitivity Test Given in Table III.

## IV. THE METHODS OF DIETER AND AHRENS

During the writing of this thesis, a pre-publication chapter from a book by U. Dieter and J. H. Ahrens was received. This chapter deals with the generation of non-uniform variates and contains a section on the generation of Poisson variates. Some of the methods which they described will be outlined here. They have not been programmed or tested against the composite method of this thesis.

### A. UNIT MEAN METHODS

It is well-known that the sum of independent Poisson deviates with mean $\mu_1, \mu_2,,,$ is Poisson distributed with mean $\mu_1+\mu_2+...$ Hence the following algorithm may be considered.

1. Split the mean m of the Poisson distribution into $m' \leftarrow [m]$ and $f \leftarrow m-m' < 1$. (Hence [m] means take the integer part of m and f is then the fractional part.)

2. Take m' samples from the Poisson distribution of mean 1 using some suitable method. Let s be the sum of these samples.

3. Take a further sample n from the Poisson (1) distribution. If n = 0 deliver $k \leftarrow s$. Otherwise, generate $U_1, U_2 ... U_n$ and count the number, j, of $U_i$ for which $U_i \leq f$, and deliver $k \leftarrow s+j$.

The validity of step 3 is proved as follows: the probability of observing j samples $U_i$ less than or equal to f (once n - j samples greater than f) is

$$\binom{n}{j} f^j (1-f)^{n-j}.$$

Since n follows a Poisson distribution with mean 1 this yields

$$P_j = \sum_{n=j}^{\infty} \frac{e^{-1}}{n!} \binom{n}{j} f^j (1-f)^{n-j} = \sum_{n=j}^{\infty} \frac{e^{-1}}{j!(n-j)!} f^j (1-f)^{n-j}$$

$$= \frac{e^{-1} f^j}{j!} \sum_{n-j=0}^{\infty} \frac{(1-f)^{n-j}}{(n-j)!} = \frac{e^{-1} f^j}{j!} e^{1-f} = \frac{e^{-f} f^j}{j!} .$$

The unit-mean method becomes a suitable algorithm for the case when the exact method should not be used for generating Poisson random variates, that is, when the mean is large ($m \geq 200$). This method will then become slightly faster than the exact method, but still needs to call too many uniform deviates since it is basically the same as the exact method except for splitting the mean, compared to the square root transformation approximation method, the unit mean method may be uneconomical.

## B. THE CENTER TRIANGLE METHOD

This method is suitable in the case of <u>variable</u> means m of the Poisson distribution. The mean is split into $m' = [m]$, the integer part of m, and the remaining fraction $0 \leq f = m - m' < 1$. The algorithm uses a mixture of two distributions. The first distribution has an easy sampling method and is employed most of the time. It is based on the following isosceles triangle.

$$(B.1) \qquad g(x) = \begin{cases} \frac{b}{s} - \frac{b}{s^2}\,|x-m'| & \text{if } m'-s \le x \le m'+s \\ 0 & \text{elsewhere.} \end{cases}$$

The constant s is chosen as

$$(B.2) \qquad s = 2.2160358671665\,\sqrt{m'} - 0.78125.$$

The area under g(x) is b. If b is small enough then the inequalities

$$(B.3) \qquad t_i = \int_i^{i+1} g(x)dx \le P_i \quad (i = 0,1,2,\ldots)$$

may hold for all i. The maximum b for which this is true was found numerically for all $4 \le m \le 200$. The choice of the constant $\alpha = 2.2160358671665$ in (B.2) is well motivated. The vertices of the maximum isosceles triangle that can be placed wholly inside the standard normal curve are situated at $(\pm\alpha,0)$ and $(0,1/\sqrt{2\pi})$. As m' goes to infinity, $b_{max}$ approaches the area of this triangle which is .88407040229876. The correction $-0.78125 = -25/32$ in (B.2) improves the value of $b_{max}$ for small m. [7, page 11-15, table.] Here the variable mean m is split into m = m'tf as in the unit mean method. If $m \le 8$ the exact method is applied. If $m \ge 9$, the triangular method is used with the triangular probabilities $t_i = 27/32$. This center triangle method needs a long program including several tables which requires a large memory space and suffers from the roundoff errors' of b.

This method proves a very efficient algorithm based on a simple idea. However, it is somewhat messy to describe the program and has little mathematical appeal compared to the square root transformation method. This method also needs a long execution time as does the exact method. For example, to compute one Poisson variate from a distribution with mean $m = 100$, the program would require 1886 μs [7, page 11-19].

## C. THE GAMMA METHOD

In order to obtain a sample k from the Poisson distribution with variable mean m, select a positive integer or (typically n is a little smaller than m). Then, take a sample x from a Gamma distribution with parameter n.

Case (1) if $x > m$, return a sample k from the binomial distribution with parameters $n-1$, $m/x$.

Case (2) if $x \leq m$, take a sample j from the Poisson distribution of mean $m-x$ and return $k \leftarrow n+j$.

The sample x simulates the n-th event (arrival) in a Poisson process of rate 1. If $x > m$, then there are $n-1$ arrivals in the interval $(0,x)$, and each of these has a probability of $m/x$ of being below m (Case (1)). If $x \leq m$, then the n simulated arrivals are all before m and the sample j indicates the additional events between x and m (Case (2)).

A formal proof of the procedure runs as follows. In the first case one has $m < x < \infty$ and

$$P_k = \int_m^{\infty} \binom{n-1}{k} \left(\frac{m}{x}\right)^k \left(1-\frac{m}{x}\right)^{n-1-k} \frac{e^{-x}\, x^{n-1}}{\Gamma(n)}\, dx$$

is the probability of obtaining k from the binomial distribution (n-1, m/x) summed over the Gamma (n)-distributed values of $\chi$ above m. The expression transforms into

$$P_k = \int_m^{\infty} \frac{(n-1)!}{k!(n-1-k)!} \frac{m^k (x-m)^{n-1-k}}{x^{n-1}} \frac{e^{-x}\, x^{n-1}}{(n-1)!}\, dx$$

$$= \frac{m^k}{k!(n-1-k)!} \int_m^{\infty} (x-m)^{n-1-k}\, e^{-x} dx$$

$$= \frac{m^k\, e^{-m}}{k!(n-1-k)!} \int_0^{\infty} t^{n-1-k}\, e^{-t}\, dt$$

where t = x - m. Notice that $\Gamma(n-k)$ is (n-1-k)!. Hence

$$P_k = \frac{m^k\, e^{-m}}{k!}$$

as required. In the <u>second</u> case one has $0 \leq x \leq m$, and

$$P_k = \int_0^{m} \frac{(m-x)^{k-n}\, e^{-(m-x)}}{(k-n)!} \frac{e^{-x}\, x^{n-1}}{\Gamma(n)}\, dx$$

is the probability of obtaining j = k-n from the Poisson distribution of mean m-x summed over the Gamma (n)-distributed values of x between 0 and m. This expression transforms into

$$P_k = \frac{e^{-m}}{(k-n)!\,(n-1)!} \int_0^m x^{n-1} (m-x)^{k-n} \, dx.$$

Introducing $t = x/m$ yields

$$P_k = \frac{e^{-m} m^{n-1} m^{k-n} m}{(n-1)!(k-n)!} \int_0^1 t^{n-1} (1-t)^{k-n} \, dt.$$

The integral is a Beta function $\beta(n,k-n+1)$

$$P_k = \frac{e^{-m} m^k}{(n-1)!\,(k-n)!} \; \frac{\Gamma(n)\,\Gamma(k-n+1)}{\Gamma(k+1)} = \frac{e^{-m} m^k}{k!}$$

as before. The following algorithm is considered.

1. Initialize $k \leftarrow 0$, $w \leftarrow m$

2. If $w \geq C$ ($C = 24$ was used as the mean cut-off point) go to 6

3. (Start Case (2)). Set $p \leftarrow 1$ and calculate $b \leftarrow e^{-w}$

4. Generate U and set $p \leftarrow pU$. If $p<b$ deliver k

5. Increase $k \leftarrow k+1$ and go to 4

6. Set $n \leftarrow [dw]$ where $d = 7/8$. Take a sample x from the Gamma-(n) distribution. If $x>w$ go to 8.

7. Set k k+w, w w-x and go to 2

8. (Start Case (1)). Set $p \leftarrow w/x$

9. Generate U. If $U \leq p$ increase $k \leftarrow k+1$

10. Set $n \leftarrow n-1$. If $n>1$ go to 9

11. Deliver k.

The performance of the algorithm depends on the cut-off point C in Step 2. Step 3, Step 4 and Step 5 are exactly the same as the steps of exact methods. This gamma method has a

complex sub algorithm for the binomial distribution in Steps 8-10 and requires memory space for the gamma distribution. This method may be efficient in cases with extremely large means, say $m = 1000$, but this method requires a rather long execution time as does the exact method. For example, to generate a sample from a Poisson distribution with mean 100 takes 1175 μs, see [7, page 11-23].

## LIST OF REFERENCES


1. Cox, D. R. and Lewis, P. A. W., The Statistical Analysis of Series of Events, London, Methuen, 1966.

2. Lewis, P. A. W., Goodman, A. S. and Miller, J. M., Pseudo-Random Number Generator for the System/360, IBM Systems Journal, No. 2, 1969.

3. Learmonth, G. P. and Lewis, P. A. W., Naval Postgraduate School Random Number Generator Package LLRANDOM, research report NPS 55 LW 73061A, Naval Postgraduate School, Monterey, California, June 1973.

4. Anscombe, F. J., The Transformation of Poisson Binomial and Negative Binomial Data, Biometrika, Vol. 35, December 1945.

5. Kolmogorov, A. N., Elements of the Theory of Functions and Functional Analysis Graylock Press; Academic Press, 1957-1961.

6. Lancaster, H. O., The Chi-Square Distribution, New York, Wiley, 1969.

7. Ahrens, J. H. and Dieter, U., Non-Uniform Random-Numbers, to be published.

8. Breiman, L., Statistics with a View Toward Application, Houghton Mifflin, 1973.

9. Barlett, M. S., The Square Root Transformation in the Analysis of Variance, Journal of the Royal Statistical Society, Supplement, Vol. 3, No. 68, 1936.

INITIAL DISTRIBUTION LIST

| | | No. Copies |
|---|---|---|
| 1. | Defense Documentation Center<br>Cameron Station<br>Alexandria, Virginia 22314 | 2 |
| 2. | Library, Code 0212<br>Naval Postgraduate School<br>Monterey, California 93940 | 2 |
| 3. | Chairman, Code 55<br>Department of Operations Research<br>and Administrative Sciences<br>Naval Postgraduate School<br>Monterey, California 93940 | 1 |
| 4. | Professor P. A. W. Lewis, Code 55 Le<br>Department of Operations Research<br>and Administrative Sciences<br>Naval Postgraduate School<br>Monterey, California 93940 | 1 |
| 5. | G. P. Learmonth<br>Computer Center<br>Naval Postgraduate School<br>Monterey, California 93940 | 1 |
| 6. | System Analysis Group<br>Department of Operations Naval Headquarter<br>Republic of Korea Navy<br>Seoul | 1 |
| 7. | Defence Institute of Technology and Science<br>Korea, Seoul | 1 |
| 8. | Library of Korea Naval Academy<br>Chine Hae, Korea | 1 |
| 9. | C. H. Pak<br>169-28 Sinsa Dong Sudaemunku<br>Seoul, Korea | 1 |
| 10. | U. C. Choe<br>Georgia Institute of Technology<br>225 North Avenue, N.W.<br>Atlanta, Georgia 30332 | 1 |